AD-A120 369 BROWN UNIV PROVIDENCE RI LEFSCHETZ CENTER FOR DYNAMI--ETC F/G 12/1
ON DIFFUSION APPROXIMATION OF CONTROLLED QUEUEING PROCESSES.(U)
JUN 82 Y LIAO AFOSR-81-0116
UNCLASSIFIED LCDS-82-18 AFOSR-TR-82-0885 NL

END
DATE
FILMED
11-82

AFOSR-TR- 82-0885

2

AD A120369

Lefschetz Center for Dynamical Systems

DTIC
ELECTE
S OCT 1 8 1982 D
B

DISTRIBUTION STATEMENT A
Approved for public release;
Distribution Unlimited

Division of Applied Mathematics

Brown University Providence RI 02912

ON DIFFUSION APPROXIMATION

OF CONTROLLED QUEUEING PROCESSES

by

Yu-Chung Liao

June 1982 LCDS #82-18

ON DIFFUSION APPROXIMATION

OF CONTROLLED QUEUEING PROCESSES*

by

Yu-Chung Liao
Division of Applied Mathematics
Lefschetz Center for Dynamical Systems
Brown University
Providence, Rhode Island 02912

June 1982

* This research has been supported in part by the Air Force Office of Scientific Research under Contract #AF-AFOSR 81-0116 and in part by the National Science Foundation under Contract #MCS 79-03554.

On diffusion approximation of controlled queueing processes

by

Yu-Chung Liao

Abstract

~~Consider~~ a queueing system can be controlled by switching service rate. When there is a cost to change service rate, the control problem turns out to be a sequential decision problem, i.e., to find a sequence of optimal stopping times to switch service rate. Under heavy traffic conditions, we show that the optimal cost functions of controlled rescaled queueing processes converge to that of corresponding controlled diffusions for finite time and for infinite time with discount factor criterions.

Key words: queue, heavy traffic, optimal stopping time, diffusion

1. Introduction

It is well known that heavy traffic queueing processes can be rescaled to approximate diffusion processes. Rath [3] proves that the rescaled queueing processes converge weakly to diffusion processes under certain stationary control strategies. In this paper, we consider a queueing system controlled by a finite set of control actions. Each action represents a service rate of the server of the queue. There are holding costs depending on queue length, operating costs to utilize the server and switching costs to switch control action. We show that the optimal cost functions of the controlled rescaled queueing processes converge to that of controlled diffusion processes. Hence, as in Kushner [3], it is meaningful to use diffusion processes to model queueing systems from a control point of view.

Some weak convergence properties of queueing processes are discussed in the next section. We define the control problem in Section 3 and prove the convergence of the optimal cost functions for finite time and for infinite time with discount factor criterions.

2. Weak Convergence of Queues

We follow Iglehart-Whitt [2] and Rath [4] to state the assumptions needed for weak convergence. Let $\{u(m,n),\ m \geq 1\}$ and $\{v_j(m,n),\ m \geq 1\}$ be sequences of independent and identically distributed random variables for each $j \in A = \{1,2,\ldots M\}$ and $n \geq 1$. For each $n \geq 1$, each of the sequences is independent of the other. $\{u(m,n)\}$ denotes the interarrival times of n-th queueing system and $\{v_j(m,n)\}$ denotes the service times of n-th queueing

system when service rate j (i.e., control action j) is being used. We further assume that there is an $\alpha > 2$ such that

$$Ev_j(1,n)^\alpha + Eu(1,n)^\alpha < \infty < \infty \qquad n \geq 1 \text{ and } j \in A,$$

$$0 < Eu(1,n) = 1/\eta(n) \qquad n \geq 1,$$

$$0 < Ev_j(1,n) = 1/\mu_j(n) \qquad n \geq 1 \text{ and } j \in A,$$

$$0 < \lim_{n\to\infty} \eta(n) = \eta < \infty,$$

$$0 < \lim_{n\to\infty} \mu_j(n) = \mu_j < \infty \qquad j \in A,$$

$$0 < \lim_{n\to\infty} \sigma^2[u(1,n)] = \sigma_a^2 < \infty,$$

$$0 < \lim_{n\to\infty} \sigma^2[v_j(1,n)] = \sigma_j^2 < \infty \qquad j \in A.$$

$$-\infty < \lim_{n\to\infty} [\eta(n)-\mu_j(n)]n^{1/2} = d_j < \infty \quad j \in A$$

and

$$a_j = [\eta^3\sigma_a^2 + \mu_j^3\sigma_j^2]^{1/2}.$$

In general a GI/G/1 queue is not a Markov process. But we can make it Markovian by adding supplementary variables. So the state of such a process is (x,y,z) where x is the queue length, y is the elapsed time since the last customer entered the queue and z is the elapsed time since the beginning of current service and $z = 0$ if $x = 0$. Let

$$\overline{Q}^{\,n,j}_{(x,y,z)}(t) = (\overline{x}_n(t), \overline{y}_n(t), \overline{z}_n(t))$$

be the n-th queueing process under control action j with initial state

(x,y,z). Define the normalized process as

$$Q^{n,j}_{(x,y,z)}(t) = (x_n(t), y_n(t), z_n(t))$$

$$= \left(\frac{1}{\sqrt{n}}\,\bar{x}_n(nt), \frac{1}{n}\,\bar{y}_n(nt), \frac{1}{n}\,\bar{z}_n(nt)\right)$$

where

$$(\bar{x}_n(t), \bar{y}_n(t), \bar{z}_n(t)) = \bar{Q}^{n,j}_{(\sqrt{n}\,x, ny, nz)}(t).$$

As $n \to \infty$, $Q^{n,j}$ converges to a process

(1) $$Q^{0,j}(t) = (x_0(t), y_0(t), z_0(t)).$$

Here $x_0(t)$ is a Brownian motion on R^+ with drift d_j, variance a_j^2 and reflected instantaneously at origin and $y_0(t) = z_0(t) = 0$ for all t. More precisely, let all the processes above be defined on a probability space (Ω,F,P).

<u>Theorem 1.</u> <u>Let</u> (x_n, y_n, z_n) <u>be the initial state of</u> $Q^{n,j}$, $\delta(n) = |x_n - x_0|$ <u>and</u> $T < \infty$. <u>Then there is a sequence</u> $q'(n)$ <u>independent of initial states and</u> j <u>such that</u>

(2) $$P\{\sup_{0\le t\le T} |x_n(t) - x_0(t)| > q'(n) + \delta(n)\} < q'(n)$$

<u>and</u> $q'(n) \to 0$ <u>as</u> $n \to \infty$.

<u>Proof.</u> The probability measures of $Q^{n,j}$ converge weakly to that of $Q^{0,j}$ if $\delta(n) \to 0$ as $n \to \infty$ is proved in Iglehart-Whitt [2]. In (2) the fact that we can use sup-norm rather than Skorokhod metric is due to the Holder continuity of Brownian motion.

3. Convergence of Cost Functions

Let the queueing processes be controlled by switching service rate. At the instant of switch if there is a customer being served by service rate i, then the incompleted service will be discarded, i.e., the service life time will become zero, and the customer will be served by another service rate immediately. The control problem is a sequential decision problem with a deterministic impulse of the z-component of the state at the time of switch. From an optimal stopping time point of view the "small impulse" means a slight variation of obstacle. To make notations manageable, we will restrict to the M/M/1 case by assuming $u(1,n)$ and $v_j(1,n)$ are exponentially distributed for all n and j. Thus, the queueing processes are Markovian and no supplementary variable is needed. Since we are going to assume the cost functions depend on queue length only, the argument for GI/G/1 case is the same. Let

$$S_0 = R^+,$$

$$S_n = \{k/n^{1/2} | k \text{ is a nonnegative integer}\} \qquad n \geq 1,$$

$$\Omega_n = \{w_n : R^+ \to S_n | w_n \text{ is right continuous with left limit}\} \qquad n \geq 0,$$

$$x_n(t)(w_n) = w_n(t) \qquad w_n \in \Omega_n \quad \text{and} \quad n \geq 0,$$

$$F_t^n = \sigma\{x_n(s) | s \leq t\} \qquad n \geq 0,$$

$$F^n = \bigvee_{t=0}^{\infty} F_t^n \qquad n \geq 0$$

and

$$\theta_t^n : \Omega_n \to \Omega_n$$

be defined by

$$\theta^n_t(w_n)(s) = w_n(s+t) \qquad n \geq 0 \text{ and } t \geq 0.$$

Then we use

$$(3) \qquad (\Omega_n, F^n_t, F^n, \theta^n_t, x_n(t), P^n_{x_n,j})$$

to denote the *n-th* rescaled queueing process for $n > 0$ and the Brownian motion in (1) for $n = 0$. Let

$$S = \bigtimes_{n=0}^{\infty} S_n,$$

$$\Omega = \bigtimes_{n=0}^{\infty} \Omega_n,$$

$$F_t = \bigtimes_{n=0}^{\infty} F^n_t,$$

$$X(t) = \bigtimes_{n=0}^{\infty} x_n(t),$$

$$P_{X,j} = \bigtimes_{n=0}^{\infty} P^n_{x_n,j}$$

where

$$X = (x_0, x_1, x_2, \ldots) \in S \text{ and}$$

$$\theta_t(w)(s) = w(s+t) \qquad w \in \Omega \text{ and } t \geq 0.$$

Then

$$(\Omega, F_t, F, \theta_t, X(t), P_{X,j})$$

is a Markov process for each j. Its *n-th* projection, $x_n(t)$, is the process

in (3). Define $u = \{s(k),u(k)\}_{k=1}^{\infty}$ as an admissible strategy if

$$(4)\qquad \begin{array}{l} s(k) \text{ is an } F_t\text{-stopping time such that} \\ 0 \leq s(k) \leq s(k+1) \text{ and } u(k) \text{ is an A-valued} \\ F_{s(k)}\text{-measurable random variable such that} \\ u(k) \neq u(k+1) \text{ for all } k \geq 1. \end{array}$$

Also, let $U = \{u|u$ is an admissible strategy$\}$. By Robin [6], for each $j \in A$, $X \in S$ and $u \in U$ there is a unique sequence of probability measures $\{P_{X,j}^{u,k}\}_{k=0}^{\infty}$ on Ω such that

$$(5)\qquad \begin{array}{l} P_{X,j}^{u,0} = P_{X,j}, \\ P_{X,j}^{u,k} = P_{X,j}^{u,k-1} \quad \text{on } F_{s(k)} \end{array}$$

and

$$(6)\qquad P_{X,j}^{u,k}(B \cap \theta_{s(k)}^{-1}B') = E_{X,j}^{u,k-1}[I_B P_{X(s(k)),u(k)}(B')]$$

where $B' \in F$, $B \in F_{s(k)}$, $B \subset \{s(k) < \infty\}$ and I_B is the indicator function of B. From now on, when a real number x is considered as an element of S_n we mean x represents

$$\max\{y \in S_n | y \leq x\}$$

and when x is considered as an element of S we mean

$$x = (x,x,x,\ldots) \in S.$$

Lemma 1. *Given* $x \in R^+$, $j \in A$, $u = \{s(k),u(k)\}_{k=1}^{\infty}$ *and* $T < \infty$ *there is a sequence* $q(n)$ *depending on* T *only such that*

$$P_{x,j}^{u,k}\{\sup_{0\le t\le T}|x_n(t)-x_0(t)| > (k+1)q(n)\} < (k+1)q(n) \tag{7}$$

and $q(n) \to 0$ *as* $n \to \infty$.

Proof. Let $q'(n)$ be as in Theorem 1 and

$$q(n) = q'(n) + n^{-1/2}.$$

Then (7) holds for $k = 0$ by (2) and (5). Let

$$B = \{\sup_{0\le t\le T\wedge s(1)}|x_n(t)-x_0(t)| < q(n)\}.$$

Then

$$P_{x,j}^{u,1}(B \cap \{\sup_{s(1)\wedge T\le t\le T}|x_n(t)-x_0(t)| > 2q(n)\}$$

$$\le E_{x,j}^{u,o}\{I_B P_{X(s(1)\wedge T)),u(1)}(\sup_{0\le t\le T}|x_n(t)-x_0(t)| > 2q(n))\}$$

$$\le q(n).$$

Hence (7) holds for $k = 1$. Given (7) for k we can prove it for $k + 1$ in the same manner.

Let the holding and operating cost be described by $f:R^+\times A \to R^+$ such that

$$|f(x,j) - f(y,j)| < L|x-y| \quad j \in A \text{ and } x,y \in R^+$$

and

$$f(x,j) < K \quad j \in A \text{ and } x \in R^+$$

for some constant K and L. Let $C: A \times A \to R$ be the switching cost such that

$$C(i,j) > C > 0 \qquad i \neq j$$

and

$$C(i,i) = 0 \qquad i \in A$$

for some constant C. Given $r \geq 0$ as discount factor and $0 < T \leq \infty$ as terminal time, define the cost functions of initial state (x,j) and strategy u as

$$(8) \qquad J_n^m(x,j,u,T) = E_{x,j}^{u,m-1} \int_0^{T \wedge s(m)} e^{-rt} f(x_n(t),u(t))dt + \sum_{k=1}^{m} I_{\{s(k)<T\}} e^{-rs(k)} D(u(k-1),u(k)),$$

$$J_n(x,j,u,T) = \lim_{m\to\infty} J_n^m(x,j,u,T),$$

$$(9) \qquad \tilde{J}_n^m(x,j,u,T) = E_{x,j}^{u,m} \int_0^T e^{-rt} f(x_n(t),u(t))dt + \sum_{k=1}^{m} I_{\{s(k)<T\}} e^{-rs(k)} C(u(k-1),u(k))$$

where $C(u(0), u(1)) = C(j,u(1))$ and

$$u(t) = \begin{cases} j & t \in [0,s(1)), \\ u(k) & t \in [s(k),s(k+1)), \\ u(m) & t > s(m) \end{cases}$$

for all $n \geq 0$ and $m > 0$. Then

$$V_n(x,j,T) = \inf_u J_n(x,j,u,T)$$

is the optimal cost function to control the n-th rescaled queueing process for $n > 0$ and that of reflected Brownian motion for $n = 0$ and

$$V_n^m(x,j,T) = \inf_u J_n^m(x,j,u,T)$$

is that with no more than m switches.

Lemma 2. If $T < \infty$, then

(10) $$V_n^m(x,j,T) - V_n(x,j,T) = 0(1) \qquad n \geq 0.$$

Proof. Consider $m > 1$ and $u \in U$ such that

$$J_n(x,j,u,T) \leq TK$$

and

$$P_{x,j}^{u,m}(s(m) < T) = a$$

then

$$a \leq TK/C(m-1).$$

Hence, we have

$$\begin{aligned} &\tilde{J}_n^m(x,j,u,T) - J_n(x,j,u,T) \\ &\leq E_{x,j}^{u,m} \int_{T\wedge s(m)}^{T} f(x_n(t),u(t))dt \\ &\leq T^2K^2/C(m-1). \end{aligned}$$

This implies (10).

<u>Lemma 3.</u> <u>Given</u> $T < \infty$ <u>and</u> $q(n)$ <u>as in Lemma</u> 1, <u>then there is a function</u> $g(m)$ <u>such that</u>

$$|V_n^m(x,j,T) - V_0^m(x,j,T)| \leq g(m)q(n) \tag{11}$$

<u>for all</u> $n \geq 0$, $x \in R^+$ <u>and</u> $j \in A$.

<u>Proof.</u> For any $u \in U$ we have

$$\begin{aligned} &|\tilde{J}_n^m(x,j,u,T) - \tilde{J}_0^m(x,j,u,T)| \\ &\leq E_{x,j}^{u,m} \int_0^T e^{-rt} |f(x_n(t),u(t)) - f(x_0(t),u(t))| dt \\ &\leq KT(m+1)q(n) + TL(m+1)q(n) \\ &\equiv g(m)q(n). \end{aligned}$$

This implies (11).

<u>Theorem 2.</u> <u>If</u> $T < \infty$ <u>then</u>

$$V_n(x,j,T) \to V_0(x,j,T) \tag{12}$$

<u>uniformly in</u> x <u>as</u> $n \to \infty$ <u>for all</u> $j \in A$.

<u>Proof.</u> By Lemmas 2 and 3 we have

$$\begin{aligned} &|V_n(x,j,T) - V_0(x,j,T)| \\ &\leq |V_n(x,j,T) - V_n^m(x,j,T)| + |V_n^m(x,j,T) - V_0^m(x,j,T)| + |V_0^m(x,j,T) \\ &\qquad - V_0(x,j,T)| \\ &\leq q(n)g(m) + 0(1) \end{aligned}$$

for all $x \in R^+$ and $j \in A$.

Corollary 1. If $r > 0$ then (12) holds for $T = \infty$.

Remark 1. Let $T = \infty$, $n \geq 0$, $r > 0$,

$$\overline{V}_n^0(x,j,T) = E\int_0^\infty e^{-rt}f(x_n(t),j)dt$$

and

$$\overline{V}_n^m(x,j,T) = \inf_s E\int_0^s e^{-rt}f(x_n(t),j)dt + e^{-rs}\{\min_{i\neq j} C(j,i) + \overline{V}_n^{m-1}(x_n(s),i,T)\} \qquad m \geq 1. \tag{13}$$

Here s ranges over all F_t^n-stopping times and the expectations are taken with respect to $P_{x,j}$. Let

$$\overline{V}_n(x,j,T) = \lim_{m\to\infty} \overline{V}_n^m(x,j,T).$$

By Robin [6], $\overline{V}_n$ is the optimal cost function to control n-th system with admissible strategy in (4) adapted to $\{F_t^n\}$ and system with admissible strategy in (4) adapted to F_t^n and

$$V_n^m(x,j,T) = \inf_s E\int_0^s e^{-rt}f(x_n(t),j)dt + e^{-rs}\{\min_{i\neq j} C(j,i) + V_n^{m-1}(x_n(s),i,T)\} \tag{14}$$

where s ranges over all F_t-stopping times for all $m \geq 1$ and $n \geq 0$. It is clear that

$$\overline{V}_n^0(x,j,T) = V_n^0(x,j,T).$$

By Dynkin [1], the optimal stopping time in (14) depends on the state of X_n, hence, adapts to F_t^n. From (13) and (14), we have

$$\overline{V}_n^m(x,j,T) = V_n^m(x,j,T) \qquad m \geq 1 \text{ and } n \geq 0$$

by induction on m. Hence,

$$\overline{V}_n(x,j,T) = V_n(x,j,T) \qquad n \geq 0.$$

So we did not change the problem by considering a larger class of admissible controls in (4). The same is true for $T < \infty$.

Remark 2. In GI/G/1 case, Remark 1 can be repeated if $Q^{n,j}$ is Fellerien for all n and j. See Dynkin [1] and Robin [6].

Remark 3. The above argument can be generalized to the multi-dimensional case since we did not use 1-dimension property. See Riemann [5].

Acknowledgement. I wish to thank Professor Wendell H. Fleming for his encouragement, advice and careful reading of my thesis of which this is a part.

REFERENCES

[1] E. Dynkin, Optimal selection of stopping time for a Markov process, Soviet Math., 4(1963), 627-629

[2] D. Iglehart and W. Whitt, Multiple channel queues in heavy traffic, Adv. Appl. Prob., 2(1970), 150-177.

[3] H. Kushner, Probability Methods for Approximation in Stochastic Control and for Elliptic Equation, Academic Press, 1977.

[4] J. Rath, Controlled queues in heavy traffic, Adv. Appl. Prob., 7(1975), 656-671.

[5] M. Rieman, Queueing networks in heavy traffic, TR.No. 76, Dept. of Operations Research, Stanford U., 1977.

[6] M. Robin, Impulsive control of Markov processes, Ph.D. thesis, Paris U. IX, 1978.

dpl

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER AFOSR-TR- 82-0885	2. GOVT ACCESSION NO. AD-A120369	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) ON DIFFUSION APPROXIMATION OF CONTROLLED QUEUEING PROCESSES		5. TYPE OF REPORT & PERIOD COVERED TECHNICAL
		6. PERFORMING ORG. REPORT NUMBER LCDS #82-18
7. AUTHOR(s) Yu-Chung Liao		8. CONTRACT OR GRANT NUMBER(s) AFOSR-81-0116
9. PERFORMING ORGANIZATION NAME AND ADDRESS Lefschetz Center for Dynamical Systems, Division of Applied Mathematics, Brown University, Providence RI 02912		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS PE61102F; 2304/A4
11. CONTROLLING OFFICE NAME AND ADDRESS Directorate of Mathematical & Information Sciences Air Force Office of Scientific Research Bolling AFB DC 20332		12. REPORT DATE June 1982
		13. NUMBER OF PAGES 13
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)		15. SECURITY CLASS. (of this report) UNCLASSIFIED
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION ST. 1ENT (of t· abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY TES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Queue; heavy traffic; optimal stopping time; diffusion.

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

Consider a queueing system can be controlled by switching service rate. When there is a cost to change service rate, the control problem turns out to be a sequential decision problem, i.e., to find a sequence of optimal stopping times to switch service rate. Under heavy traffic conditions, the author shows that the optimal cost functions of controlled rescaled queueing processes converge to that of ccrresponding controlled diffusions for finite time and for infinite time with discount factor criterions.

DD FORM 1 JAN 73 1473

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

ATE

MED